Data-Driven Control under Input and Measurement Noise

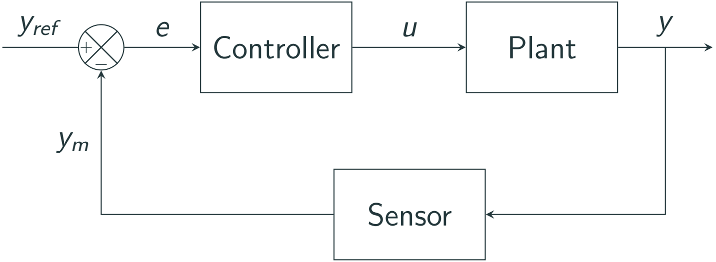
Jared Miller
Tianyu Dai
Mario Sznaier

Feb. 21, 2023
NYU MERIIT Lab Seminar Series

Northeastern University

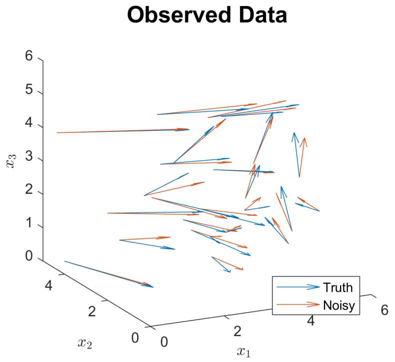
What is Data-Driven Control?

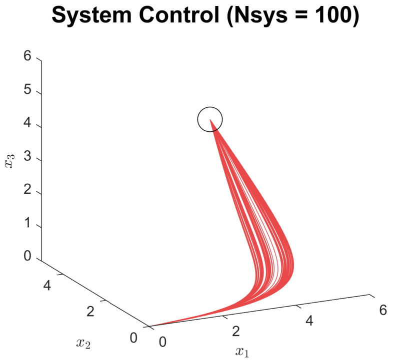
Design a controller for an unknown plant



Control system directly from data, no sysid required

Example of Data-Driven Control





Single controller stabilizes all data-consistent plants

Algorithms for Data-Driven Control

Virtual Reference Feedback Tuning (first methods)

Set-Membership (this talk)

- (Data-consistent plants) $\subseteq$ (K-Stabilized plants)
- Certificates of set containment (Farkas, S-Lemma, SOS)

Behavioral

- Parameterize and pick out best system trajctory (MPC)
- Willem's Fundamental Lemma (DeePC)

Flow of Presentation

Input/Measurement noise description and challenges

Solution using polynomial optimization (superstability)

Extend to other problems (Quadratic, H2, ARX)

Noise Model and Difficulty

Error-in-Variable Noise Task

Noisy measurements $\mathcal{D} = \{\hat{x}_t, \hat{u}_t\}_{t=1}^{T}$ of linear system

$$x_{t+1} = Ax_t + Bu_t$$

Data $\mathcal{D}$ corrupted by (L_∞-bounded):

Δx : state-measurement noise

Δu : input noise

w : process noise

Find state-feedback $u = Kx$ to stabilize all plants (A, B) consistent with $\mathcal{D}$

Error-in-Variable Relations

Noise processes $\forall t = 1..T$

$$\epsilon_x \ \geq \|\Delta x_t\|_\infty \qquad \epsilon_u \geq \|\Delta u_t\|_\infty \qquad \epsilon_w \geq \|w_t\|_\infty$$

Relations $\forall t = 1..T-1$

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + Ew_t \\ \hat{x}_t &= x_t + \Delta x_t \\ \hat{u}_t &= u_t + \Delta u_t \end{aligned}$$

$(A, B, \Delta x, \Delta u, w)$ unknown, $E \in \mathbb{R}^{n \times e}$ known

Bilinear Trouble

$(A, B, \Delta x, \Delta u, w)$ all unknown

Total of $n(n+m) + T(n+m+e)$ variables

$$\hat{x}_{t+1} - \Delta x_{t+1} = A\hat{x}_t - A\Delta x_t + Bu_t - B\Delta u_t - Ew_t$$

Multiplication between unknown $A\Delta x_t$, also in $B\Delta u_t$

Stabilization task is immediately NP-hard

Even sysid is NP-hard

(probably why no one else has done EIV data-driven control)

Main Ideas

Use superstability to form a more tractable control problem

Formulate a large-scale polynomial optimization problem

Improve scalability by applying a Theorem of Alternatives

Superstability

Superstablity Definition

Superstability (Polyak 2001), $\|x\|_\infty$ is a CLF

$$\|A + BK\|_\infty < 1$$

Poles of $A + BK$ in unit diamond $\{z \mid \mathrm{Re}(z) + \mathrm{Im}(z) < 1\}$

If $\|A + BK\|_\infty = \gamma$, then $\|x_t\|_\infty \leq \gamma^{(t+1)/n}\|x_0\|_\infty$

Constant K must superstabilize all consistent (A, B)

Superstability Formulations

Linear constraints to impose superstability

Sign-based formulation, $n2^n$ linear constraints

$$\sum_{s\in\{-1,1\}^n} s_j(A+BK)_{ij} < 1 \qquad \forall i$$

Equivalent Convex Lift, $2n^2 + n$ linear constraints

$$\begin{aligned}
&\exists M \in \mathbb{R}^{n\times n}: \\
&\quad \textstyle\sum_{j=1}^{m} M_{ij} < 1 && \forall i \\
&\quad -M_{ij} \leq (A+BK)_{ij} \leq M_{ij} && \forall i,j
\end{aligned}$$

Full Program

Consistency Set

Consistency set $\bar{\mathcal{P}}(A, B, \Delta x)$ (with $\epsilon_u = \epsilon_w = 0$)

$$\bar{\mathcal{P}}: \left\{ \begin{aligned} &0 = -\Delta x_{t+1} + A\Delta x_t + h_t^0 && \forall t = 1..T-1 \\ &\|\Delta x_t\|_\infty \leq \epsilon_x && \forall t = 1..T \end{aligned} \right\}$$

Affine weight h^0 is defined by,

$$h_t^0 = \hat{x}_{t+1} - A\hat{x}_t - Bu_t \qquad \forall t = 1..T-1.$$

Assumption: enough data collected such that $\bar{\mathcal{P}}$ compact

Superstability for Plants

Set of plants consistent with $\mathcal{D}$ (with projection π):

$$\mathcal{P}(A, B) = \pi^{A,B}\bar{\mathcal{P}}(A, B, \Delta x)$$

Find $K \in \mathbb{R}^{m\times n}$ such that $(A + BK)$ is Schur $\forall (A, B) \in \mathcal{P}$

Restrict to superstability: $\|A + BK\|_\infty < 1, \quad \forall (A, B) \in \mathcal{P}$

Superstability Application

Superstability certificate $M(A, B) : \mathcal{P} \to \mathbb{R}^{n \times n}$

$2n^2 + n$ inequality expressions over $\bar{\mathcal{P}}$ (margin $\delta > 0$)

$$\forall i = 1..n : 1 - \delta - \sum_{j=1}^{n} M_{ij}(A, B) \geq 0 \tag{1a}$$

$$\forall i = 1..n,\ j = 1..n : \tag{1b}$$

$$M_{ij}(A, B) - (A_{ij} + \sum_{\ell=1}^{m} B_{i\ell} K_{\ell j}) \geq 0$$

$$M_{ij}(A, B) + (A_{ij} + \sum_{\ell=1}^{m} B_{i\ell} K_{\ell j}) \geq 0$$

LP in (M, K) for each $(A, B) \in \mathcal{P}$ (infinite dimensional)

Can choose M to be continuous in compact $\mathcal{P}$

Sum-of-Squares Method

Nonnegative $q(x) \in \mathbb{R}[x]$ is SOS ($q \in \Sigma[x]$) if there exists a vector $v(x) \in \mathbb{R}[x]^s$, *Gram* matrix $Z \in \mathbb{S}_+^s$ with $q = v^T Z v$

Putinar Positivestellensatz (Psatz) nonnegativity certificate over set $\mathbb{K} = \{x \mid g_i(x) \geq 0, h_j(x) = 0\}$:

$$q(x) = \sigma_0(x) + \textstyle\sum_i \sigma_i(x) g_i(x) + \sum_j \phi_j(x) h_j(x)$$
$$\exists \sigma_0(x) \in \Sigma[x], \quad \sigma_i(x) \in \Sigma[x], \quad \phi_j \in \mathbb{R}[x].$$

Psatz at degree $2d$ is an SDP, monomial basis: $s = \binom{n+d}{d}$

Computational Complexity (Full)

Restrict $M_{ij}(A, B)$ to a polynomial of degree $2d$

Each infinite-dimensional linear constraint becomes an SOS constraint (Psatz) in $(A, B, \Delta x)$

Each Psatz has a PSD Gram matrix of size $\binom{n(n+m+T)+d}{d}$

$(n = 2, m = 2, T = 15, d = 2)$: size 780

Alternatives

Motivation and Size Comparison

Use Δx-affine structure of $\bar{\mathcal{P}}$ to eliminate Δx

Maximal size of Gram (PSD) matrices

Size	Full	Alternatives
Super	$\binom{n(n+m+T)+d}{d}$	$\binom{n(n+m)+d}{d}$

When $(n = 2, m = 2, T = 15, d = 2)$:

$$\text{Full} = 780, \text{ Altern.} = 45$$

Robust Counterpart Method (eliminating noise)

Polytope-constrained noise Δx

$$\Delta x \in \bar{\mathcal{P}} = \{\Delta x \mid G\Delta x \leq h,\ C\Delta x = f\}$$

All (q, G, h, C, f) are functions of (A, B) $\bar{\mathcal{P}}$

Linear inequality involving Δx

$$q(A, B) \geq 0 \qquad \forall (A, B, \Delta x) \in \bar{\mathcal{P}}$$

Equivalent (nonconservative) Robust Counterpart without Δx

$$\exists \zeta \geq 0, \mu \mid q \geq h^T \zeta + f^T \mu,\ 0 = G^T \zeta + C^T \mu.$$

Theorem of Alternatives

Superstability condition q: Full program in $(A, B, \Delta x)$

$$q(A, B) \geq 0 \qquad \forall (A, B, \Delta x) \in \bar{\mathcal{P}}$$

Alternatives program in (A, B) with no conservatism

$$\begin{aligned}
\text{find} \quad & \zeta^{\pm}_{1:T}(A, B) \geq 0, \;\; \mu_{1:T-1}(A, B) \\
& q \geq \textstyle\sum_{t,i} \epsilon_x (\zeta^{+}_{t,i} + \zeta^{-}_{t,i}) + \sum_{t=1}^{T-1} \mu_t^T h_t^0 \quad \forall (A, B) \\
& \zeta^{+}_{1} - \zeta^{-}_{1} = A^T \mu_1 \\
& \zeta^{+}_{T} - \zeta^{-}_{T} = -\mu_{T-1} \\
& \zeta^{+}_{t} - \zeta^{-}_{t} = A^T \mu_t - \mu_{t-1} \qquad \forall t \in 2..T-1
\end{aligned}$$

Polynomial Alternatives Certificate

Choose $\zeta^{\pm}$ SOS, μ polynomial when $\bar{\mathcal{P}}$ compact

Express SOS Alternatives certificate as $q(A, B) \in \Sigma^{\mathrm{alt}}[\mathcal{P}]$

Find degree-$2d$ polynomial matrix $M_{ij}(A, B)$ with

$$
\begin{aligned}
&\forall i = 1..n : 1 - \delta - \textstyle\sum_{j=1}^{n} M_{ij}(A, B) \in \Sigma^{\mathrm{alt}}[\mathcal{P}] \\
&\forall i = 1..n,\ j = 1..n : \\
&\quad M_{ij}(A, B) - (A_{ij} + \textstyle\sum_{\ell=1}^{m} B_{i\ell} K_{\ell j}) \in \Sigma^{\mathrm{alt}}[\mathcal{P}] \\
&\quad M_{ij}(A, B) + (A_{ij} + \textstyle\sum_{\ell=1}^{m} B_{i\ell} K_{\ell j}) \in \Sigma^{\mathrm{alt}}[\mathcal{P}]
\end{aligned}
$$

$\zeta^{\pm},\ \mu$: same multiplicity as SOS Psatz multipliers over $\bar{\mathcal{P}}$

Further notes about complexity

In practice $d = 1$ suffices for Alternatives while $d = 2$ is required for Full

With $(n = 2, m = 1, d_{\mathrm{alt}} = 1, d_{\mathrm{full}} = 2)$

Maximum size PSD matrices

	Gram	ζ	μ (vector)
Alternatives	7	7	7
Full (T = 4)	120	15	120
Full (T = 6)	190	19	190
Full (T = 8)	276	23	276

All Noise

All Noise Consistency Set

Consistency set $\bar{\mathcal{P}}^{\mathrm{all}}(A, B, \Delta x, \Delta u, w)$:

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + Ew_t & \forall t = 1..T-1 \\
\hat{x}_t &= x_t + \Delta x_t, \quad \hat{u}_t = u_t + \Delta u_t & \forall t = 1..T-1 \\
\epsilon_x &\geq \|\Delta x_t\|_\infty, \ \epsilon_u \geq \|\Delta u_t\|_\infty, \ \epsilon_w \geq \|w_t\|_\infty & \forall t = 1..T
\end{aligned}
$$

Set of consistent plants,

$$
\mathcal{P}^{\mathrm{all}}(A, B) = \pi^{A,B}\bar{\mathcal{P}}^{\mathrm{all}}(A, B, \Delta x, \Delta u, w)
$$

$(\Delta x, \Delta u, w)$ together not much more complex than Δx alone

All Noise Size

Use Alternatives to eliminate $(\Delta x, \Delta u, w)$

Maximal size of Gram (PSD) matrices

Size	Full	Alternatives
Super	$\binom{n(n+m)+T(n+m+e)+d}{d}$	$\binom{n(n+m)+d}{d}$

When $(n = 2, m = 2, T = 15, d = 2, e = 1)$:

$$\text{Full} = 3570, \text{ Alternatives} = 45$$

Quadratic Stabilization

Quadratic Stabilization

Quadratic Lyapunov function $x^T Y x$ for $Y \in \mathbb{S}^n_{++}$

$$Q(A,B) = \begin{bmatrix} Y & (A+BK)Y \\ * & Y \end{bmatrix} = \begin{bmatrix} Y & AY+BS \\ * & Y \end{bmatrix} \in \mathbb{S}^{2n}_{++}$$

Recover controller $K = SY^{-1}$

Find constant (Y, K) to stabilize all $(A, B) \in \mathcal{P}$

Polynomial Matrix Inequalities

SOS method (scalar): $q(x) \geq 0$

Extend to matrices $Q(x) \in \mathbb{S}^s_{++}$

SOS matrix: $Q(x) = R(x)^T R(x) \in \Sigma^s[x]$ for matrix $R(x)$

Gram matrix (PSD) constraint of size $s\binom{n+d}{d}$

Scherer Psatz: nonnegativity over constraint sets

Quadratic Stabilization Program

Quadratic Full: Size $2n\binom{n(n+m+T)+d}{d}$

$$\begin{bmatrix} Y & AY + BS \\ * & Y \end{bmatrix} \in \Sigma^{2n}[\bar{\mathcal{P}}]_{\leq 2d} \tag{2}$$

Can eliminate Δx, form Alternatives with size $2n\binom{n(n+m)+d}{d}$

Alternatives could add conservatism

Extend to worst-case-H_2-optimal control

Single-Input Single-Output

ARX model

Autoregressive Model with Exogenous Input (ARX)

$$y_t = -\sum_{i=1}^{n_a} a_i y_{t-i} + \sum_{i=1}^{n_b} b_i u_{t-i}.$$

Data $\mathcal{D} = (\hat{u}, \hat{y})$ and no state x,

$$\begin{aligned} \hat{u} &= u + \Delta u, & \|\Delta u\|_\infty &\leq \epsilon_u \\ \hat{y} &= y + \Delta y, & \|\Delta y\|_\infty &\leq \epsilon_y \end{aligned}$$

Find controller u to stabilize (a, b) consistent with $\mathcal{D}$

Superstability for ARX

Original model with vectors (a, b)

$$y_t = -\sum_{i=1}^{n_a} a_i y_{t-i} + \sum_{i=1}^{n_b} b_i u_{t-i}.$$

Transfer Function with one-step-behind operator $\lambda u_t = u_{t-1}$

$$G(\lambda) = \frac{\sum_{i=1}^{n_b} b_i \lambda^i}{1 + \sum_{i=1}^{n_a} a_i \lambda^i} = \frac{B(\lambda)}{1 + A(\lambda)}$$

Superstability definition, linear constraints

$$\|a\|_1 < 1$$

Dynamic Compensation

Compensator $C(\lambda) = \tilde{B}(\lambda)/(1+\tilde{A}(\lambda))$

Closed-loop system

$$G_{cl}(\lambda) = \frac{G(\lambda)}{1+G(\lambda)C(\lambda)} = \frac{B(\lambda)(1+\tilde{A}(\lambda))}{(1+A(\lambda))(1+\tilde{A}(\lambda))+B(\lambda)\tilde{B}(\lambda)}.$$

Superstable: coefficients of G_{cl} denominator have L_1 norm < 1

Fixed C superstabilizes all $(A, B) \in \mathcal{P}$ (from $\mathcal{D}$)

ARX Program Sizes

Set $\mathcal{P}$ originally contains $(a, b, \Delta u, \Delta y)$

Eliminate $(\Delta u, \Delta y)$ in alternatives

Maximal size of Gram (PSD) matrices ($N = N_a + N_b$)

Size	Full	Alternatives
Super	$\binom{2N+T-1+d}{d}$	$\binom{N+d}{d}$

No conservatism in Alternatives

Superstability Examples

Example 1

Ground-truth system $n = 3, m = 2, T = 40$

$$A = \begin{bmatrix} 0.6852 & 0.0274 & 0.5587 \\ 0.2045 & 0.6705 & 0.1404 \\ 0.8781 & 0.4173 & 0.1981 \end{bmatrix}, \ B = \begin{bmatrix} 0.4170 & 0.3023 \\ 0.7203 & 0.1468 \\ 0.0001 & 0.0923 \end{bmatrix}$$

Noise parameters $\epsilon_x = 0.05, \epsilon_u = 0, \ \epsilon_w = 0$

Solve $\gamma^* = \min_{\gamma \in \mathbb{R}} \gamma : \|A + BK\|_\infty \leq \gamma$ for all $(A, B) \in \mathcal{P}$

Example 1: Complexity

Data horizon $T = 6$,

	d	#scalar variables
Full	2	3.4×10^7
Altern.	1	67776

Altern recovers ground truth $\gamma^* = 0.7259$ when $\epsilon_x = 0$

Example 1: Results

With $T = 40$:

$$\begin{aligned} \gamma^*_{\text{alt}} &= 0.8880 \quad \text{Alternatives with } d = 1 \text{ (worst-case)} \\ \gamma^*_{\text{clp}} &= 0.7749 \quad \text{Alternatives controller applied to ground truth} \\ \gamma^*_{\text{true}} &= 0.7259 \quad \text{Ground truth} \end{aligned}$$

Example 2: (Monte Carlo, Stabilization)

Ground truth system $(\epsilon_w, \epsilon_u = 0)$

$$A = \begin{bmatrix} 0.6863 & 0.3968 \\ 0.3456 & 1.0388 \end{bmatrix}, \quad B = \begin{bmatrix} 0.4170 & 0.0001 \\ 0.7203 & 0.3023 \end{bmatrix}$$

S = number of successful designs out of 100 trials

S vs. ϵ_x with $T = 8$

ϵ_x	0.05	0.08	0.11	0.14
S	100	84	57	39

S vs. T with $\epsilon_x = 0.14$

T	8	10	12	14
S	39	60	75	86

Example 3: (Monte Carlo, H2 Performance)

Median H_2 performance in 100 trials (PMI)

H_2-norm vs. ϵ_x with $T = 8$

ϵ	0.05	0.08	0.11	0.14
$\gamma_{2,\mathrm{clp}}$	1.97	2.07	2.18	2.15
$\gamma_{2,\mathrm{worst}}$	2.30	2.73	3.23	4.31

H_2-norm vs. T with $\epsilon_x = 0.14$

T	8	10	12	14
$\gamma_{2,\mathrm{clp}}$	2.07	1.96	1.94	1.93
$\gamma_{2,\mathrm{worst}}$	2.73	2.42	2.23	2.20

Example 4: (ARX Superstabilization)

Ground truth system ($\epsilon_w, \epsilon_u = 0$)

$$y_t = u_{t-2} - (0.5y_{t-1} - 1.21y_{t-2} - 0.605y_{t-3})$$

Fixed-order control $n_a = 4, n_b = 3$ with $\epsilon_y = \epsilon_u = \epsilon$

γ v.s. ϵ with $T = 80$

ϵ	0.02	0.04	0.06	0.08
γ	0.25	0.49	0.73	0.98

γ v.s. T with $\epsilon = 0.02$

T	20	40	60	80
γ	0.44	0.31	0.27	0.25

Take-aways

Conclusion

Stabilization in the Error-in-variables setting

Formulate SOS certificates over consistency set

Alternatives to simplify computational complexity

Conservatism only introduced in Quadratic Stability

Acknowledgements

Didier Henrion, POP group at LAAS-CNRS

Chateaubriand Fellowship of the Office for Science Technology of the Embassy of France in the United States.

National Science Foundation (NSF)

Air Force Office of Scientific Research (AFOSR)

Thank you for your attention

arxiv:2210.1489, 2210.14893

github:jarmill/error_in_variables

Bonus Content

Set Membership: Process Noise Alone

Superstability with only L_∞-bounded process noise (not EIV)

$$\hat{x}_{k+1} = A\hat{x}_t + B\hat{u}_t + w_t \qquad\qquad \forall t = 1..T-1$$

Polytope of data-consistent plants $P_1(A, B)$:

$$P_1 = (A, B): \ \|\hat{x}_{k+1} - A\hat{x}_t - B\hat{u}_t\|_\infty \leq \epsilon_w \quad \forall t = 1..T-1$$

Superstable-plants polytope $P_2(A, B)$ given constant (M, K)

$$P_2 = (A, B): \ -M \leq A + BK \leq M$$

Control via LP (Cheng, Sznaier, Lagoa 2015)

Sparse but Conservative Tightening

Equality constraints $0 = -\Delta x_{t+1} + A\Delta x_t + h_t^0$

Define row groups $C_i = (A_{i,1:n}, B_{i,1:m})$

Each equality constraint in (i, t) only involves one group

Sparse multipliers $\zeta_{it}^{\pm}(C_i) \geq 0,\ \mu_{it}(C_i)$

Max. Gram matrix size $\binom{n+m+d}{d}$ rather than $\binom{n(n+m)+d}{d}$

Has never worked on our experiments though